

### Warm up:

I want to look at the relationship between scores on the first test in Stat (x-variable) and scores on the second test in Stat (y-variable).

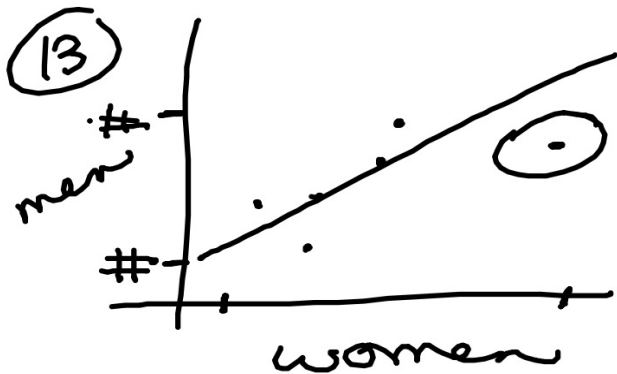
I find a correlation of  $r = -0.78$  between the two variables. Find the coefficient of determination, and interpret it.

$$r^2 = (-0.78)^2 = 0.6084 \Rightarrow 60.84\%$$

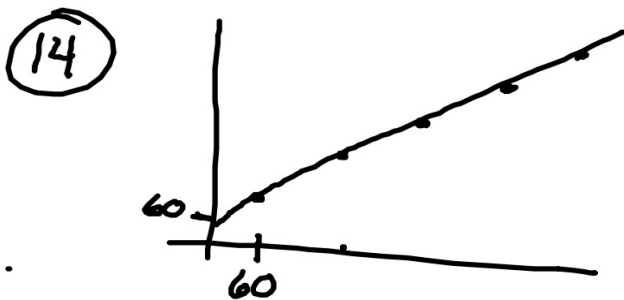
~~-0.78~~

variation

60.84% of the change in  
second test scores is due to the  
change in 1<sup>st</sup> test scores.



$r = +$   
not close to 1



$$\hat{y} = x + 3$$

27)  $\hat{y} = 1425 - 19.87x$   
 $-19.87 \text{ ft}^3$  (30)  
 1 degrees  $828.9 \text{ ft}^3$

For every 1 degree increase, the gas consump. decreases by  $19.87 \text{ ft}^3$ .

## 4.2 notes- PART 2

① -linear, negative, mod. strong

②  $\hat{y} = 17.89 - 0.587x$

$$r = -0.759$$

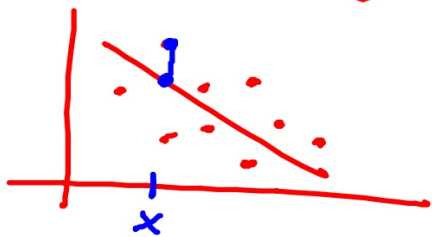
$$r^2 = 0.576$$

$$\hat{y} = 17.89 - 0.587(10)$$

$$\hat{y} = 12.016$$

## Residuals (errors)

- actual y-value - predicted y-value  
(from LSR line)

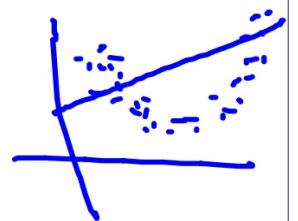
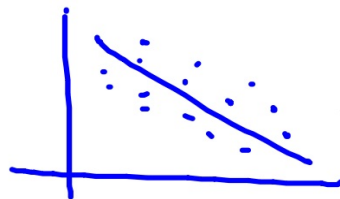


$$\sum \text{residuals} = 0$$

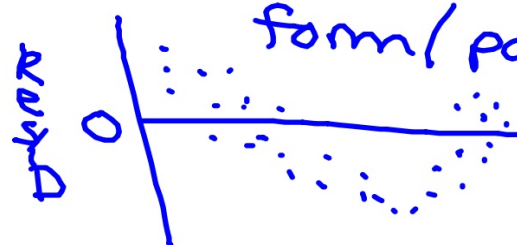
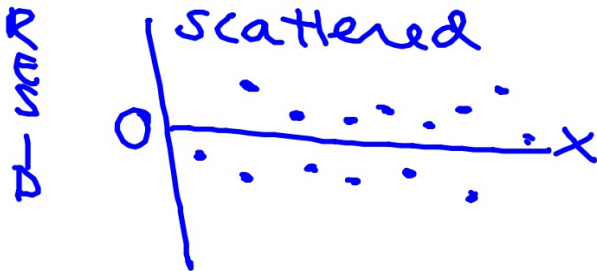
↑  
sum

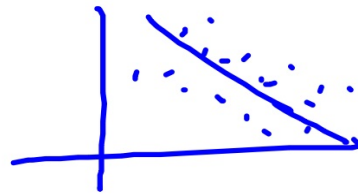
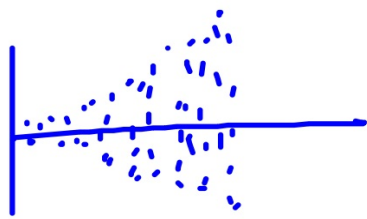
## Residual Plot

- Scatterplot of residuals (y)  
vs. x-variable



- helps assess the fit of line  
scattered form/pattern

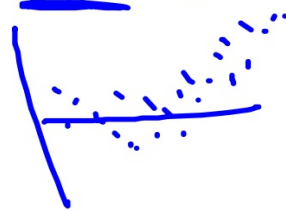
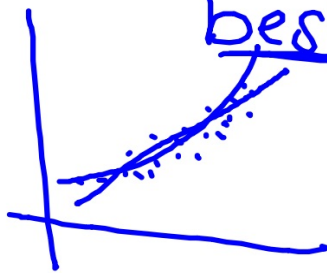




- no pattern = scattered
  - the LSR line is a good fit for the data.

- pattern =

- the LSR line is not the best fit for the data



HW:

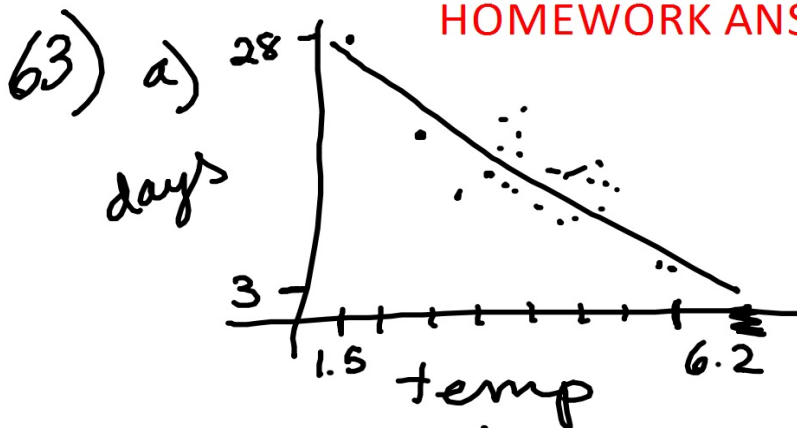
p. 198 #63

Read 4.2

Print 4.2-part 3 notes

Test Friday(?)

# HOMEWORK ANSWERS



When the temp is  $0^{\circ}\text{C}$ , the days til bloom are 33.12.

$$\text{LinReg}(a+bx) \quad X, Y, Y1$$

$$\hat{y} = 33.12 - 4.69x$$

$$\frac{y}{x} = \frac{-4.69 \text{ days}}{1^{\circ}\text{C}}$$

For every 1 degree increase the days til bloom decrease by 4.69 days.



c)  $3.5^{\circ}\text{C}$

$$\hat{y} = 33.12 - 4.69(3.5)$$

$$\hat{y} = 16.7 \text{ days}$$

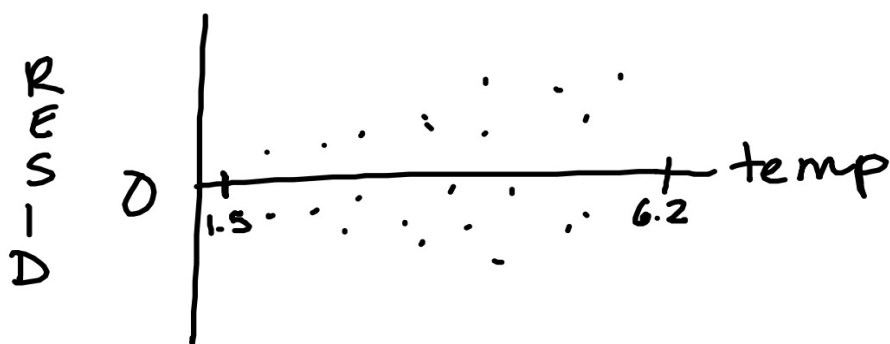
April 17<sup>th</sup>

d)  $\hat{y} = 33.12 - 4.69(4.5) = 12.015 \text{ days}$   $\checkmark^x$

resid = actual  $y$  - predicted  $y$


$$= 10 - 12.015$$

$$= -2.015 \text{ days}$$



the LSR line is a good fit  
for data, b/c the resid.  
plot is scattered.

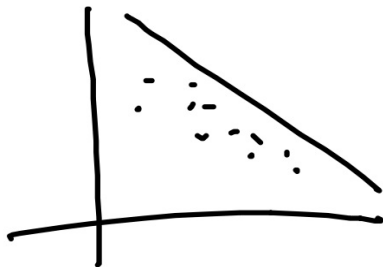
$r^2 = 72\%$  of the change in days  
is due to the change in temp.

Complete worksheet 4.2B. Answers to #8 & 9:



greater

underestimate = + resid



- resid = overestimate

⑥ \$98,957.22

HW: p. 182 #37 & 38

and finish wksht 4.2C if needed